Standardized Definitions and Approaches for Vertical Resolution and Uncertainty in the NDACC Ozone DIAL Algorithms

Thierry Leblanc [1], R. J. Sica [2], S. Godin-Beekmann [3], J. A. E. van Gijsel [4], G. Liberti [5], G. Payen [6], F. Gabarrot [6], T. Trickl [7], and A. Haefele [8]

(1) California Institute of Technology, Jet Propulsion Laboratory,
(2) The University of Western Ontario, London, Canada
(3) LATMOS-IPSL, CNRS-INSU, Paris, France
(4) Royal Netherlands Meteorological Institute (KNMI), Netherlands
(5) ISAC-CNR, Rome, Italy
(6) Obs. des Sci. de l'Univers de La Réunion, CNRS/Université de la Réunion, France
(7) Karlsruher Institut für Technologie, IMK-IFU, Garmisch-Partenkirchen, Germany
(8) Meteoswiss, Payerne, Switzerland

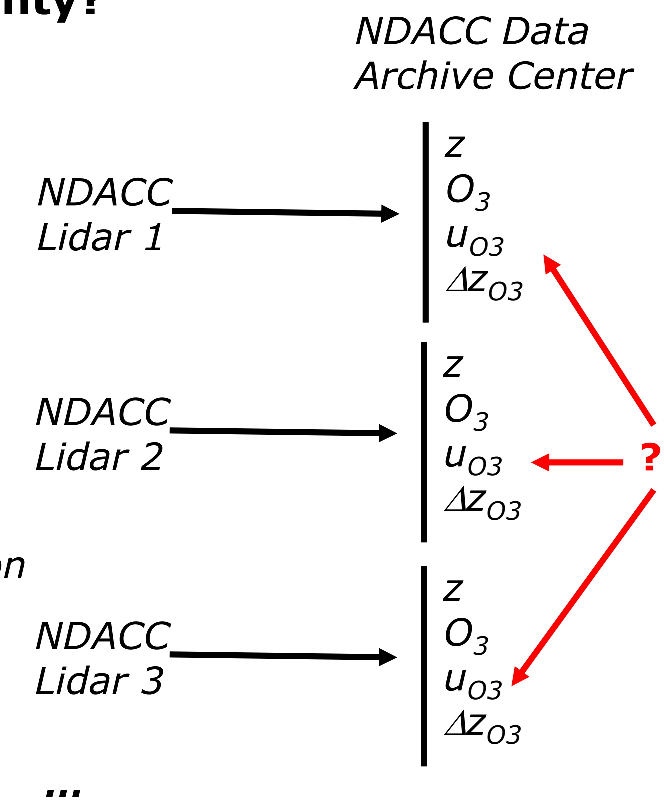
Motivation

Why standardized definitions and approaches for vertical resolution and uncertainty?

- *NDACC lidars use a wide spectrum of methodologies and technologies to measure ozone, temperature, aerosols, and water vapor*

- *As a result, it is difficult to archive measurement and analysis information consistently across the network*

- *Yet, consistent definitions are needed for data exploitation such as satellite validation, profile intercomparison, assimilation in numerical models, and trend studies*



➔ To address this, an ISSI Team composed of NDACC Lidar Working Group members recently formulated new recommendations for the use of standardized definitions and approaches leading to a network-wide, consistent reporting of vertical resolution and uncertainty in the NDACC lidar data files

Vertical Resolution

NDACC-standardized definition 1: Use FWHM of Impulse Response

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

c_n = Filter coefficient c_n at altitude $z(k)$

First, we assume that the lidar signal (or the retrieved species profile) is vertically filtered at some point during data processing

Vertical resolution

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

c_n = Filter coefficient c_n at altitude $z(k)$

$$\left[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N\right] \otimes \begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix}$$

We then convolve the filter coefficients with an Impulse…

$I_{INP}(k,n)$ = Impulse at altitude $z(k)$ and distance n from $z(k)$

Vertical resolution

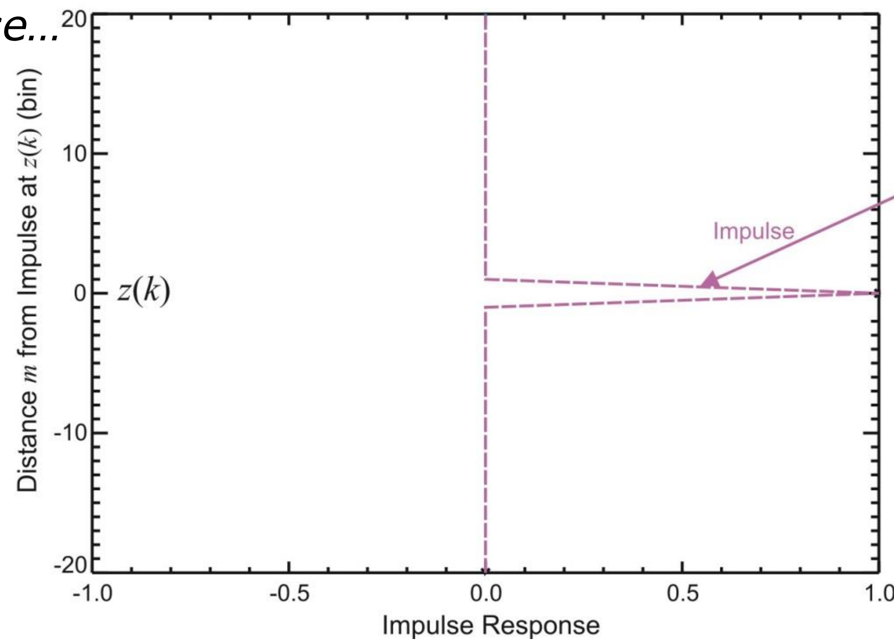
NDACC-standardized definition 1: Use FWHM of Impulse Response

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

c_n = Filter coefficient c_n at altitude $z(k)$

$$\left[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N\right] \otimes \begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix}$$

We then convolve the filter coefficients with an Impulse…



$I_{INP}(k,n)$ = Impulse at altitude $z(k)$ and distance n from $z(k)$

For smoothing filters:
Impulse = Kronecker function

For derivative filters:
Impulse = Heaviside function

Vertical resolution

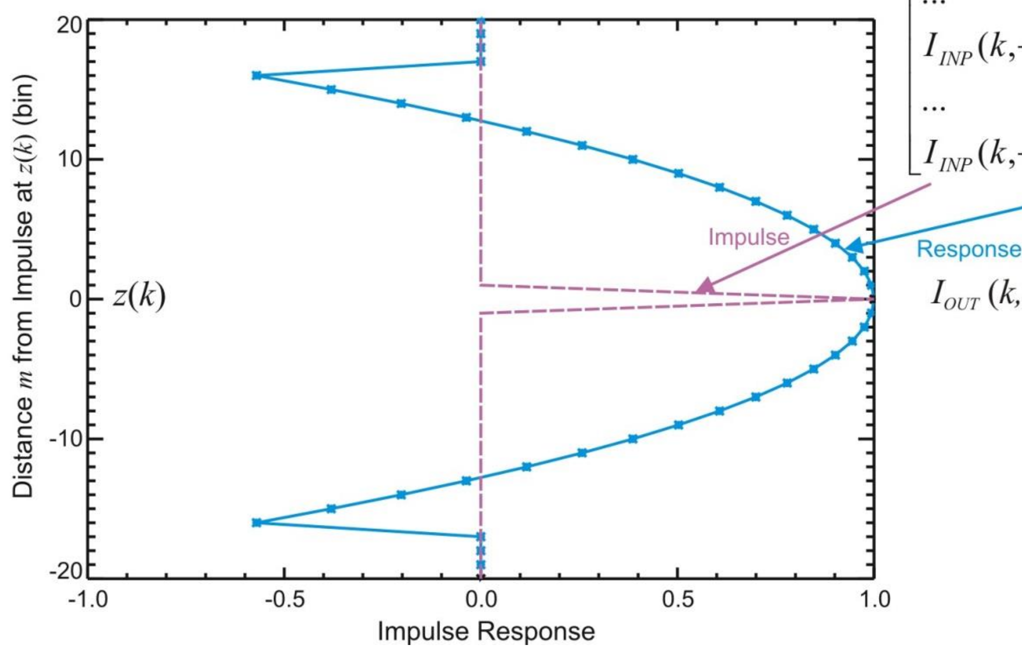
NDACC-standardized definition 1: Use FWHM of Impulse Response

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

c_n = Filter coefficient c_n at altitude $z(k)$

Compute response…

$$\left[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N\right] \otimes \begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix} = \begin{bmatrix} I_{OUT}(k,M) \\ \ldots \\ I_{OUT}(k,N) \\ \ldots \\ I_{OUT}(k,1) \\ I_{OUT}(k,0) \\ I_{OUT}(k,-1) \\ \ldots \\ I_{OUT}(k,-N) \\ \ldots \\ I_{OUT}(k,-M) \end{bmatrix}$$



$I_{OUT}(k,m)$ = Response at altitude $z(k)$ and distance m from $z(k)$

Vertical resolution

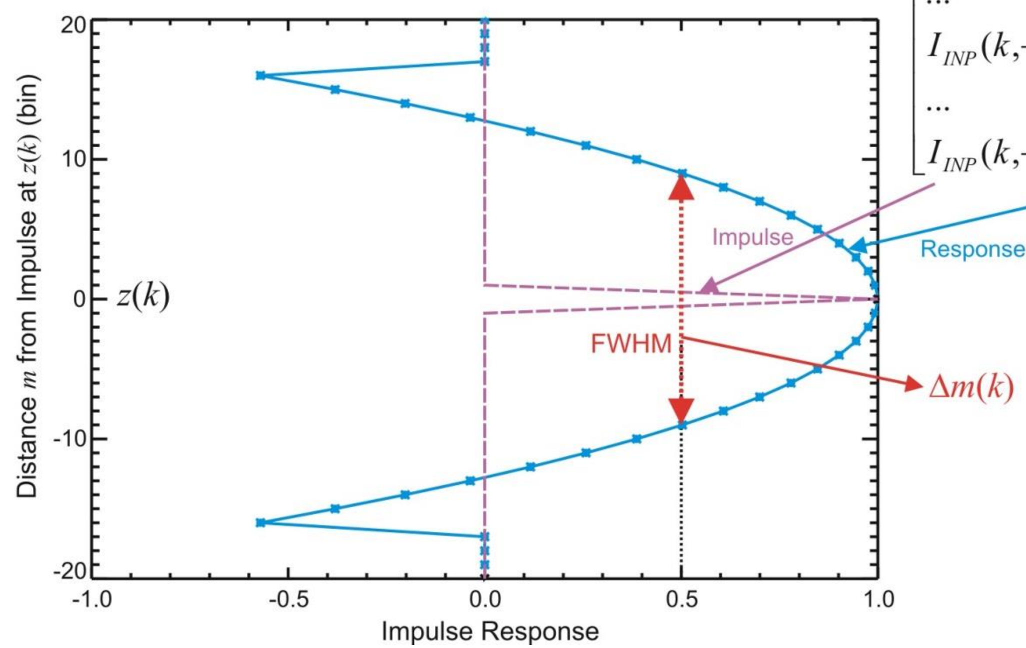
NDACC-standardized definition 1: Use FWHM of Impulse Response

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

c_n = Filter coefficient c_n at altitude $z(k)$

$$\left[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N\right] \otimes \begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix} = \begin{bmatrix} I_{OUT}(k,M) \\ \ldots \\ I_{OUT}(k,N) \\ \ldots \\ I_{OUT}(k,1) \\ I_{OUT}(k,0) \\ I_{OUT}(k,-1) \\ \ldots \\ I_{OUT}(k,-N) \\ \ldots \\ I_{OUT}(k,-M) \end{bmatrix}$$

Calculate FWHM of response…



Vertical resolution

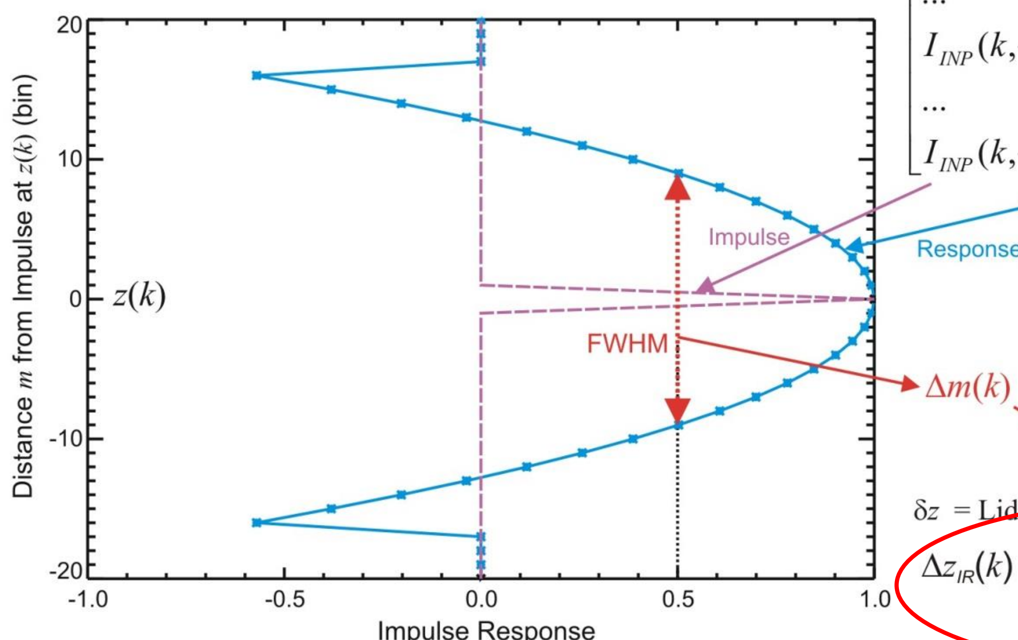
NDACC-standardized definition 1: Use FWHM of Impulse Response

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

c_n = Filter coefficient c_n at altitude $z(k)$

$$[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N] \otimes \begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix} = \begin{bmatrix} I_{OUT}(k,M) \\ \ldots \\ I_{OUT}(k,N) \\ \ldots \\ I_{OUT}(k,1) \\ I_{OUT}(k,0) \\ I_{OUT}(k,-1) \\ \ldots \\ I_{OUT}(k,-N) \\ \ldots \\ I_{OUT}(k,-M) \end{bmatrix}$$

Multiply by sampling resolution



$$\Delta z_{IR}(k) = \delta z\ \Delta m(k)$$

δz = Lidar sampling resolution

$\Delta z_{IR}(k)$ = NDACC-standardized vertical resolution at altitude $z(k)$

If multiple smoothing occurrences during data processing

Filter 1

$$[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N]\otimes\begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix} = \begin{bmatrix} I_{OUT}(k,M) \\ \ldots \\ I_{OUT}(k,N) \\ \ldots \\ I_{OUT}(k,1) \\ I_{OUT}(k,0) \\ I_{OUT}(k,-1) \\ \ldots \\ I_{OUT}(k,-N) \\ \ldots \\ I_{OUT}(k,-M) \end{bmatrix}$$

The response from the first occurrence is used as the Impulse for the second occurrence (instead of using a Kronecker or Heaviside)

Filter 2

$$[c_{-N},\ldots,c_{-1},c_0,c_1,\ldots,c_N]\otimes\begin{bmatrix} I_{INP}(k,M) \\ \ldots \\ I_{INP}(k,N) \\ \ldots \\ I_{INP}(k,1) \\ I_{INP}(k,0) \\ I_{INP}(k,-1) \\ \ldots \\ I_{INP}(k,-N) \\ \ldots \\ I_{INP}(k,-M) \end{bmatrix} = \begin{bmatrix} I_{OUT}(k,M) \\ \ldots \\ I_{OUT}(k,N) \\ \ldots \\ I_{OUT}(k,1) \\ I_{OUT}(k,0) \\ I_{OUT}(k,-1) \\ \ldots \\ I_{OUT}(k,-N) \\ \ldots \\ I_{OUT}(k,-M) \end{bmatrix}$$

And so on… until no more smoothing occurs in data processing

This propagation method is mathematically exact all the way to the final archived product!

Where is this vertical resolution reported?

Two new variables have been added to the NDACC lidar data files archived in HDF format:

The vector Δz_{IR} *just defined is reported in the following NDACC Lidar HDF variable:*
O3.NUMBER.DENSITY_ABSORPTION.DIFFERENTIAL_RESOLUTION.ALTITUDE.IMPULSE.RESPONSE.FWHM

The 2D array $I_{OUT}(nk,nm)$ *is reported in the following NDACC Lidar HDF variable:*
O3.NUMBER.DENSITY_ABSORPTION.DIFFERENTIAL_RESOLUTION.ALTITUDE.IMPULSE.RESPONSE

The vertical resolution "historically" reported by the PI is now reported in the following HDF variable:

O3.NUMBER.DENSITY_ABSORPTION.DIFFERENTIAL_RESOLUTION.ALTITUDE.ORIGINATOR

It is currently kept for consistency, but will become obsolete soon

NDACC-lidar standardized vertical resolution definition 2

Second definition (optional) based on the cut-off frequency of Digital Filters:

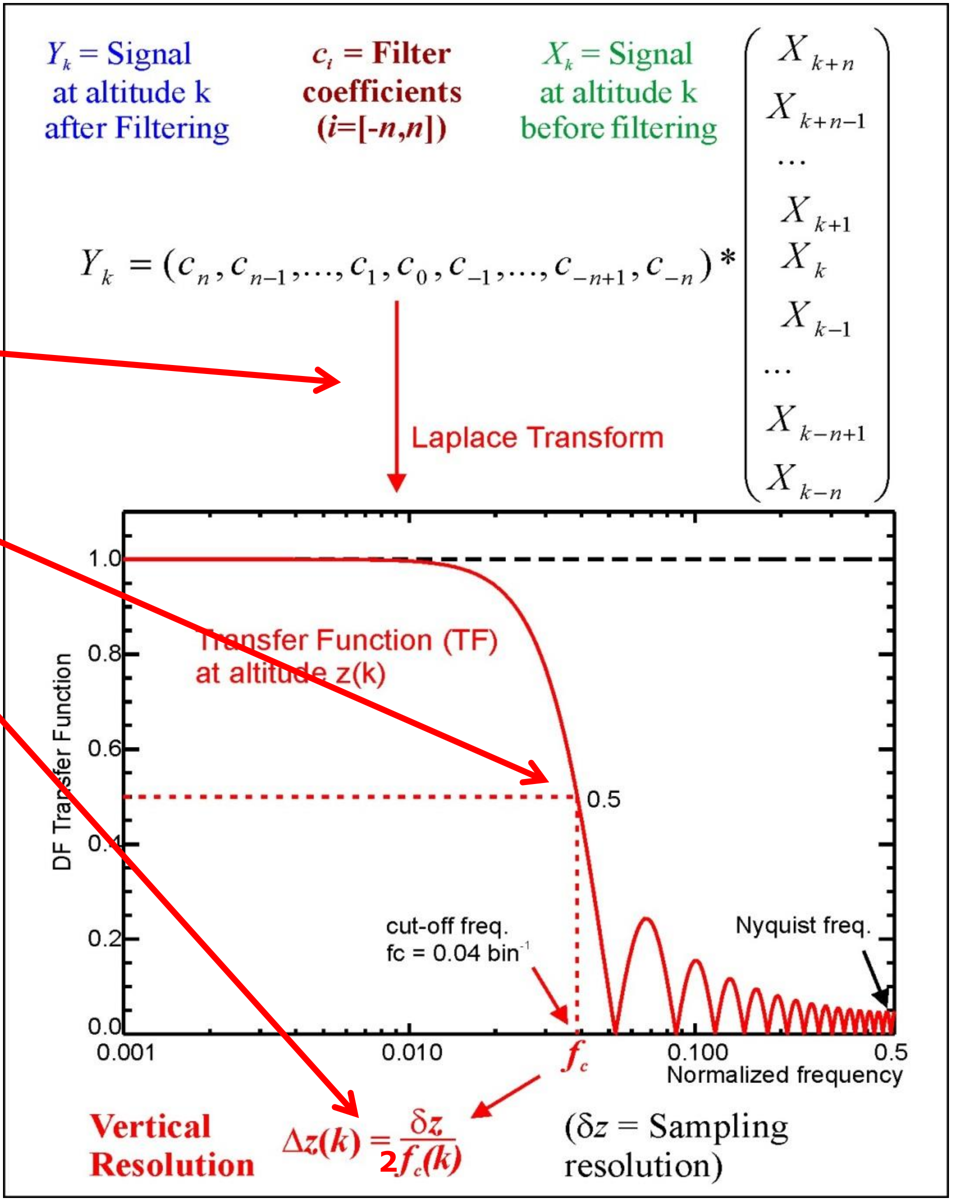
1. Apply Laplace Transform to coefs

2. Identify cut-off frequency (i.e., where Transfer Function = 0.5)

3. Take inverse of cut-off frequency and multiple by half the sampling resolution

Note the factor of "2", different from what is used in spectral analysis

This factor allows consistent values with Definition 1 (IR)



NDACC-wide (and beyond...) implementation

Numerical tools:

2 subroutines (1 per definition) written in IDL, FORTRAN, MATLAB, C++, and PYTHON, compute automatically vertical resolution following the standardized definitions, and were distributed to all NDACC lidar PIs

Documentation:

How to use the routines, and how to write the new standardized variables into the NDACC HDF data files in preparation

Note:

- *The "IR" definition (definition 1) has a physical meaning that reminds the AK reported for passive remote sensing techniques*
- *Both standardized definitions can be used likewise for water vapor and aerosol lidar*
- *Other networks, such as TOLNet (Tropospheric Ozone Lidar Network) and GRUAN (in preparation), have also adopted this standardization*

Now about NDACC-Standardized Uncertainty Budget…

Just like for vertical resolution,

it is NOT the quantitative estimates

that are being standardized,

but the definitions and approaches

Sources

For ozone DIAL, 11 independent sources suitable for standardization:

- *Detection noise*
- *Signal saturation (pile-up) correction*
- *Background noise extraction*
- *Ozone absorption cross-section*
- *Molecular extinction cross-section*
- *Ancillary air density profile (or temperature and pressure)*
- *NO_2 absorption cross-sections*
- *Ancillary NO_2 profile*
- *SO_2 absorption cross-sections*
- *Ancillary SO_2 profile*
- *O_2 absorption cross-section (Herzberg region)*

3 sources currently unsuitable for standardization:

- *Analog-to-digital signal conversion subsystems*
- *Partial beam-telescope field-of-view overlap (a.k.a "misalignment")*
- *Contamination by particulate extinction and backscatter*

➔ *Additional work is required before we can provide recommendations for a standardized treatment of uncertainty associated with these sources*

Propagation

4 main recommendations:

- *Use traceable input quantities (e.g., well-documented absorption cross-section datasets with uncertainties)*

- *Use variance propagation rule to propagate each of the 11 components in parallel, including covariance terms when necessary*

- *For fundamental physical constants: Use metrological sources (e.g., CODATA) for proper decimal truncation, and assume zero-uncertainty*

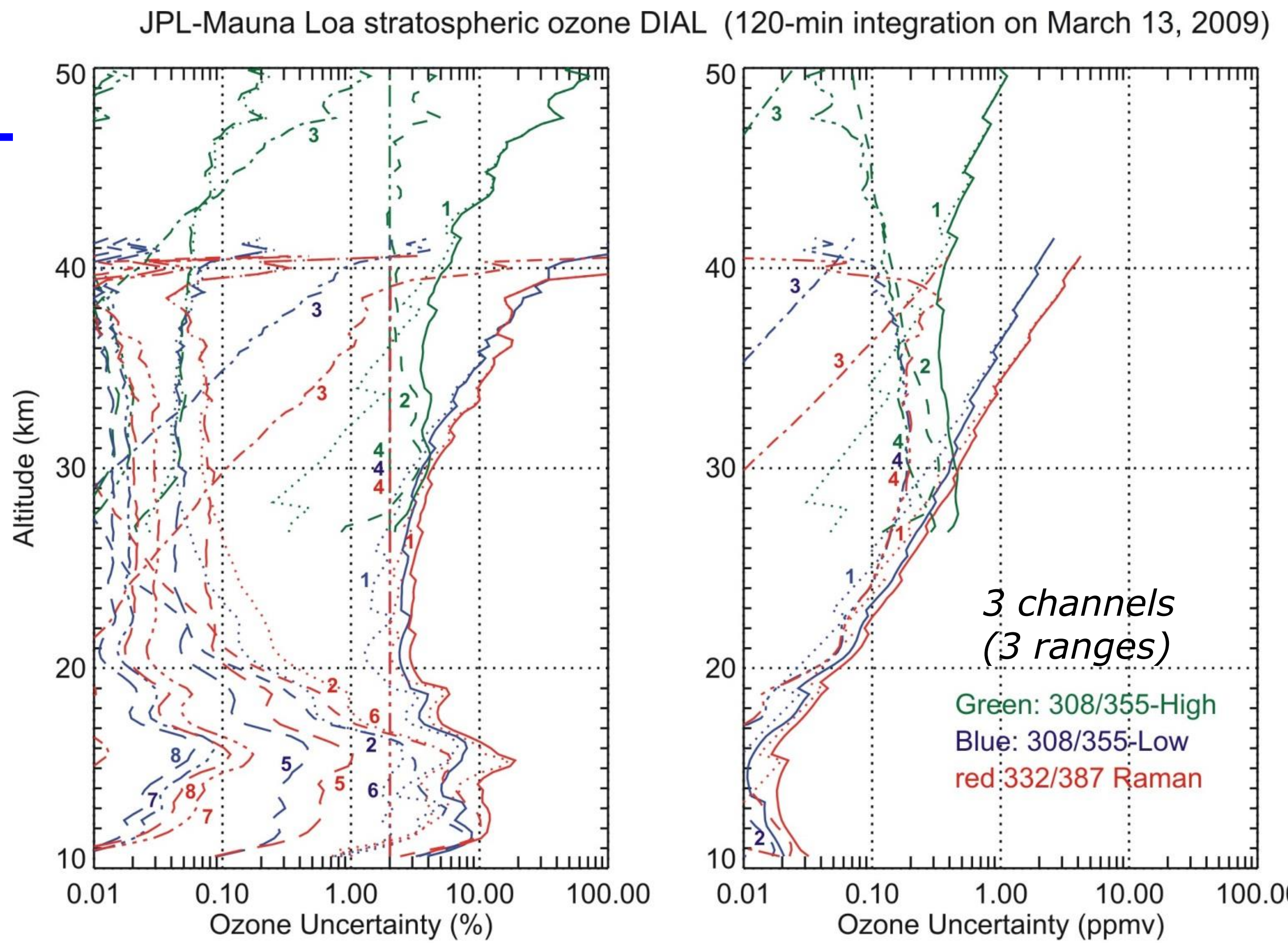
- *Combine all components only at the very end of data processing, (just before archiving in the NDACC data file)*

Quantitative example 1:

JPL stratospheric ozone lidar at Mauna Loa, Hawaii

Solid curves: Total combined uncertainty

Other curves: Individual uncertainty components

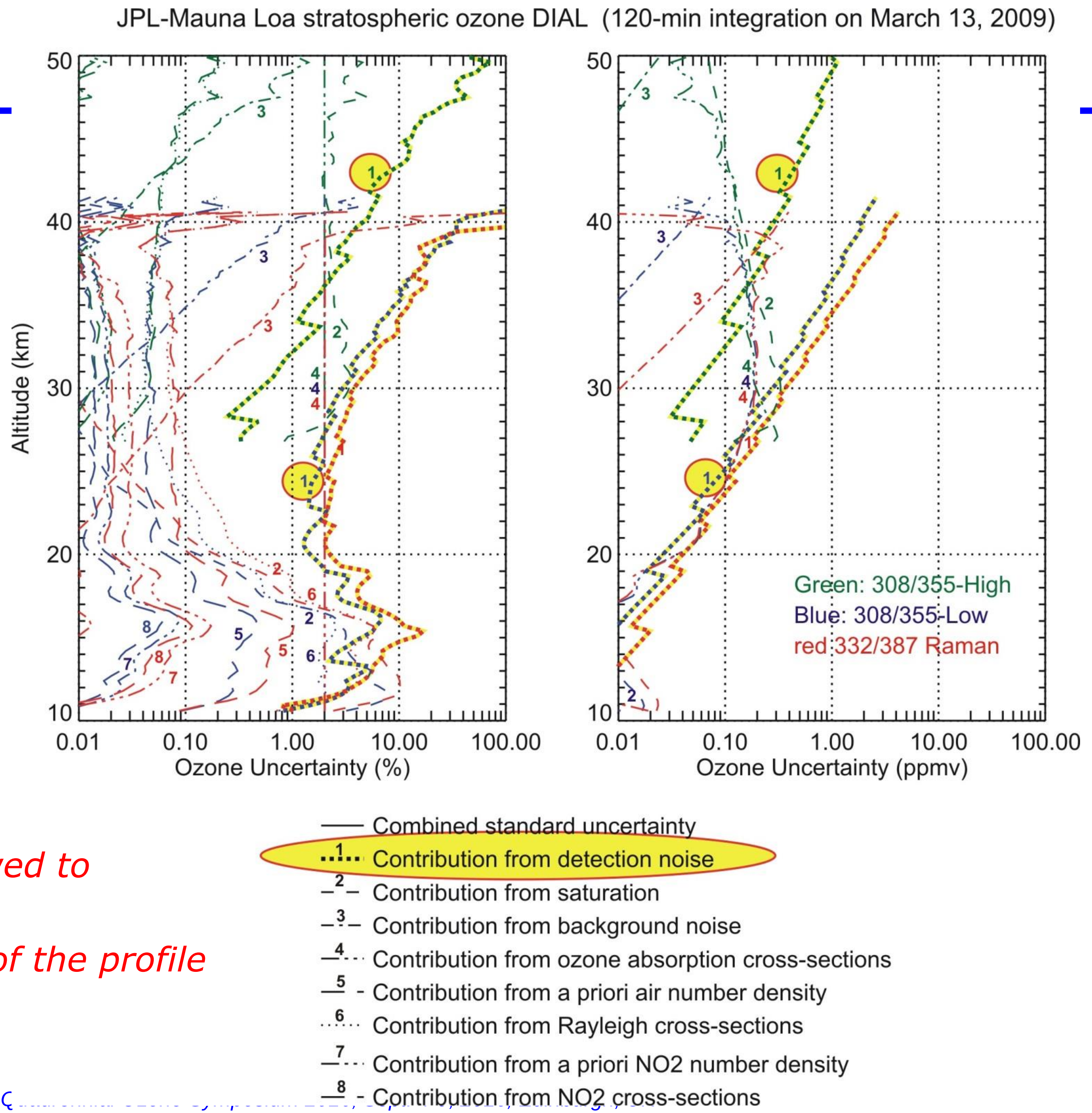


Combined standard uncertainty
1 Contribution from detection noise
2 Contribution from saturation
3 Contribution from background noise
4 Contribution from ozone absorption cross-sections
5 Contribution from a priori air number density
6 Contribution from Rayleigh cross-sections
7 Contribution from a priori NO2 number density
8 Contribution from NO2 cross-sections

Quantitative estimates can vary significantly, depending on lidar instrument considered!

Quantitative example 1:

JPL stratospheric ozone lidar at Mauna Loa, Hawaii

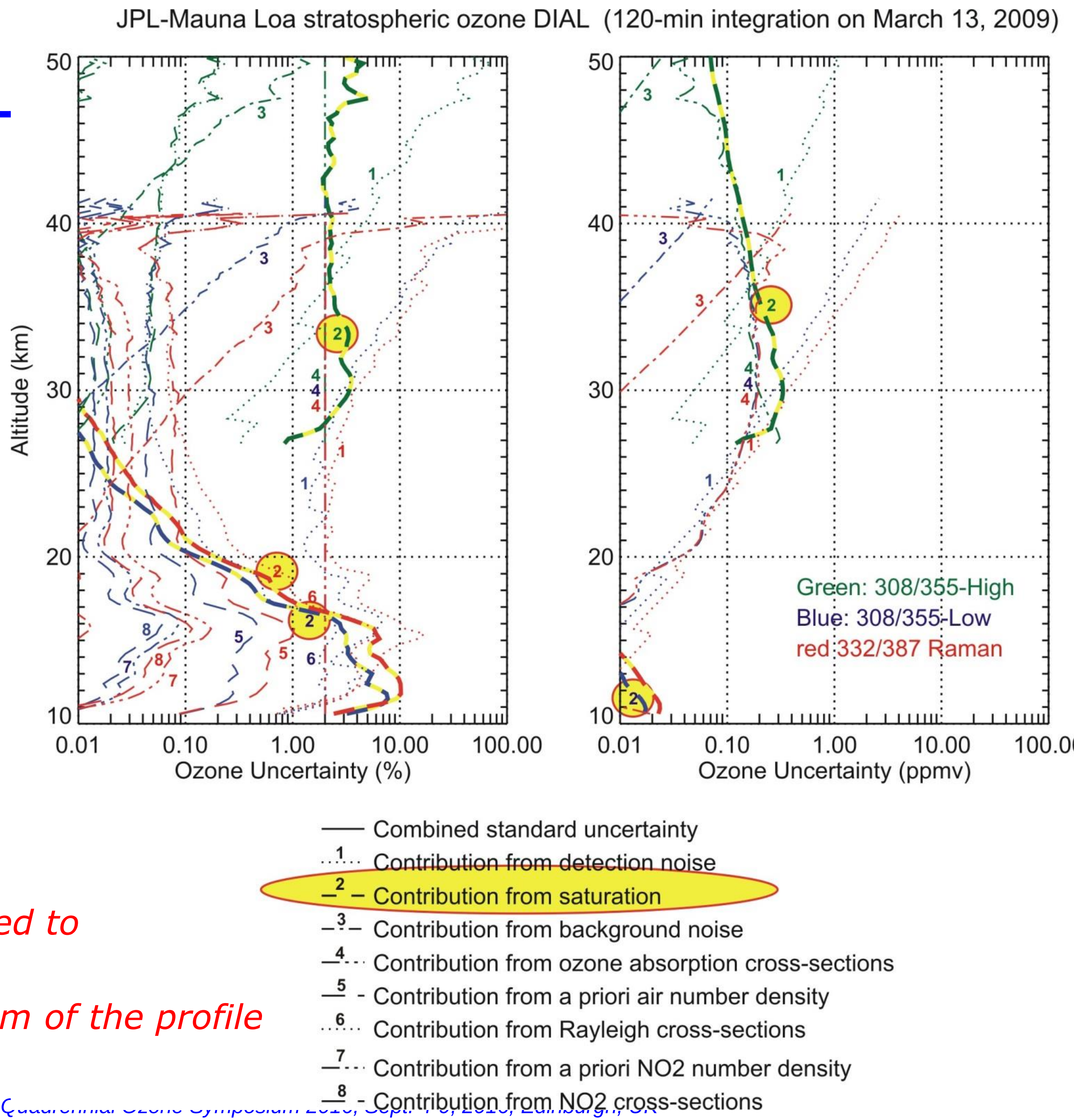


Ozone uncertainty owed to detection noise:

Dominant at the top of the profile

Quantitative example 1:

JPL stratospheric ozone lidar at Mauna Loa, Hawaii

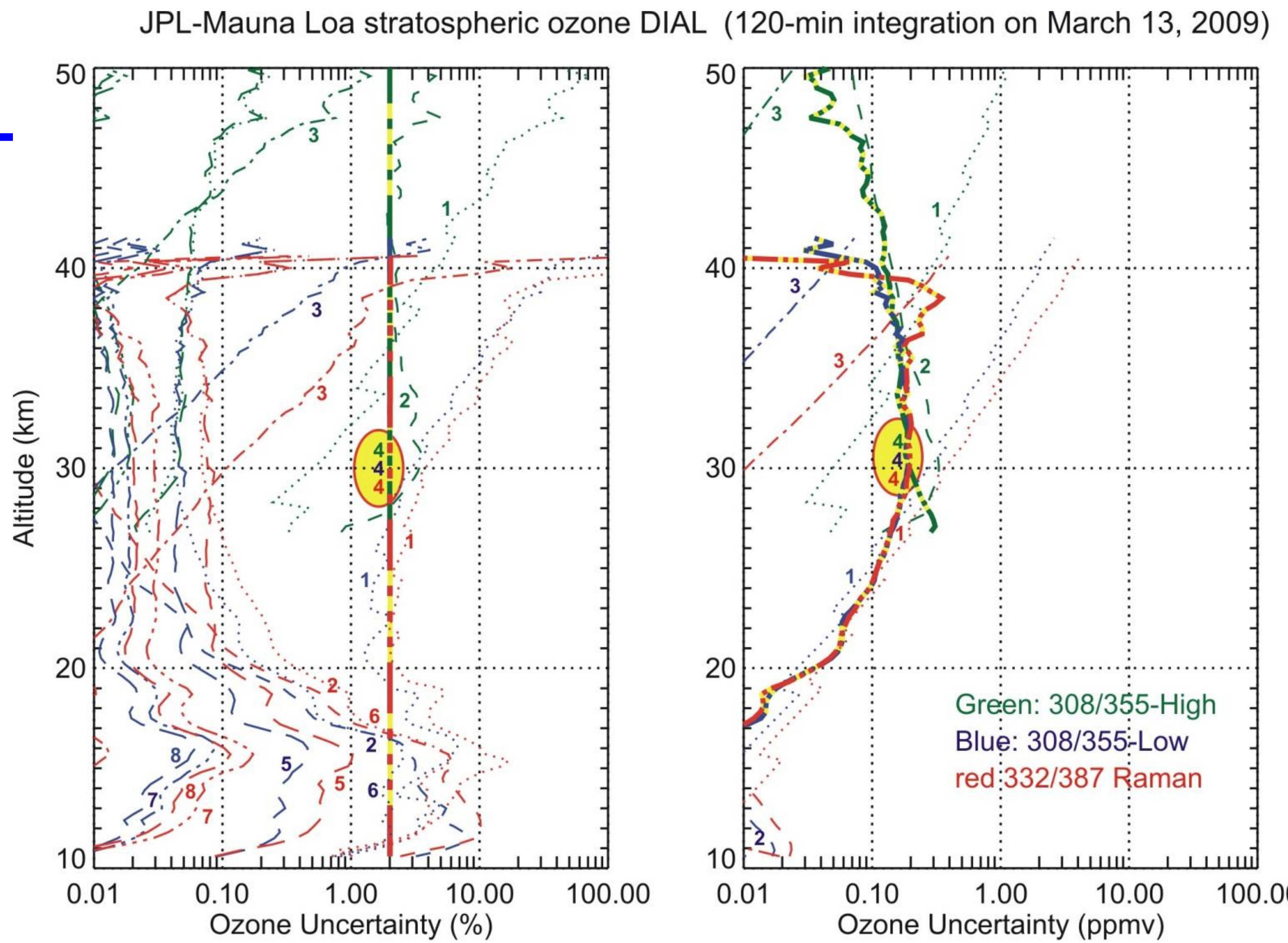


Ozone uncertainty owed to saturation correction:

Dominant at the bottom of the profile

Quantitative example 1:

JPL stratospheric ozone lidar at Mauna Loa, Hawaii



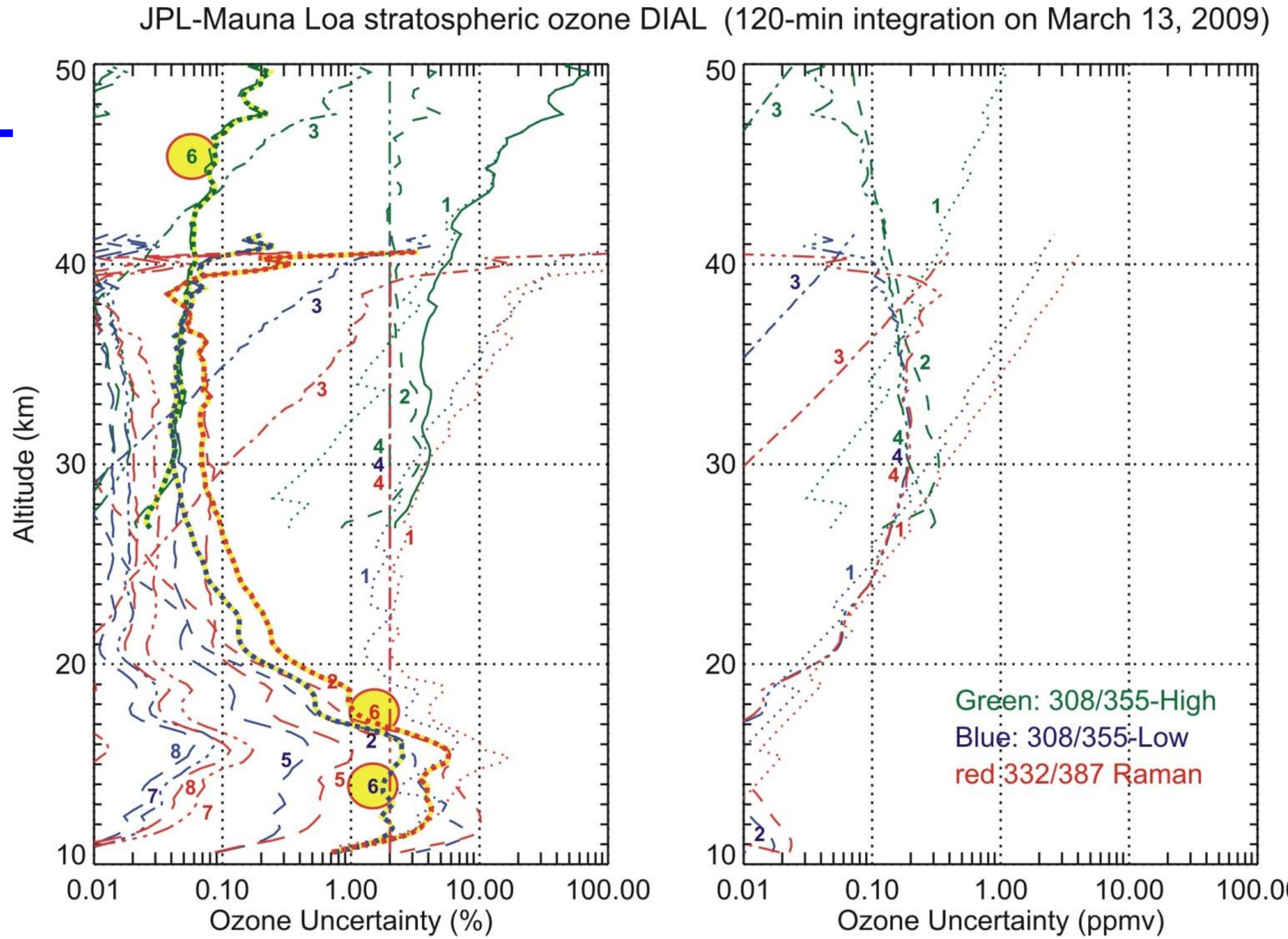
— Combined standard uncertainty
1 Contribution from detection noise
2 Contribution from saturation
3 Contribution from background noise
4 Contribution from ozone absorption cross-sections
5 Contribution from a priori air number density
6 Contribution from Rayleigh cross-sections
7 Contribution from a priori NO2 number density
8 Contribution from NO2 cross-sections

Ozone uncertainty owed to ozone absorption cross-sections:

Constant (%) throughout profile

Quantitative example 1:

JPL stratospheric ozone lidar at Mauna Loa, Hawaii



— Combined standard uncertainty
1 Contribution from detection noise
2 Contribution from saturation
3 Contribution from background noise
4 Contribution from ozone absorption cross-sections
5 Contribution from a priori air number density
6 Contribution from Rayleigh cross-sections
7 Contribution from a priori NO2 number density
8 Contribution from NO2 cross-sections

Ozone uncertainty owed to Rayleigh cross-sec. :

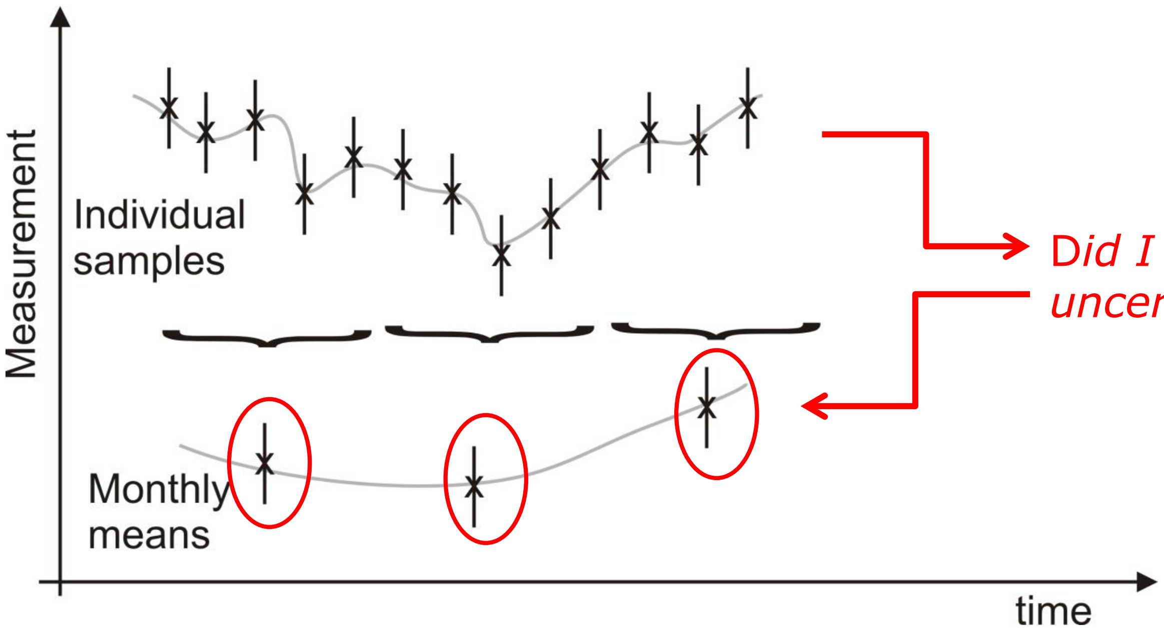
Very small except below 18 km (larger for Raman channels)

Use of Uncertainty Information by NDACC Data Users

More recommendations

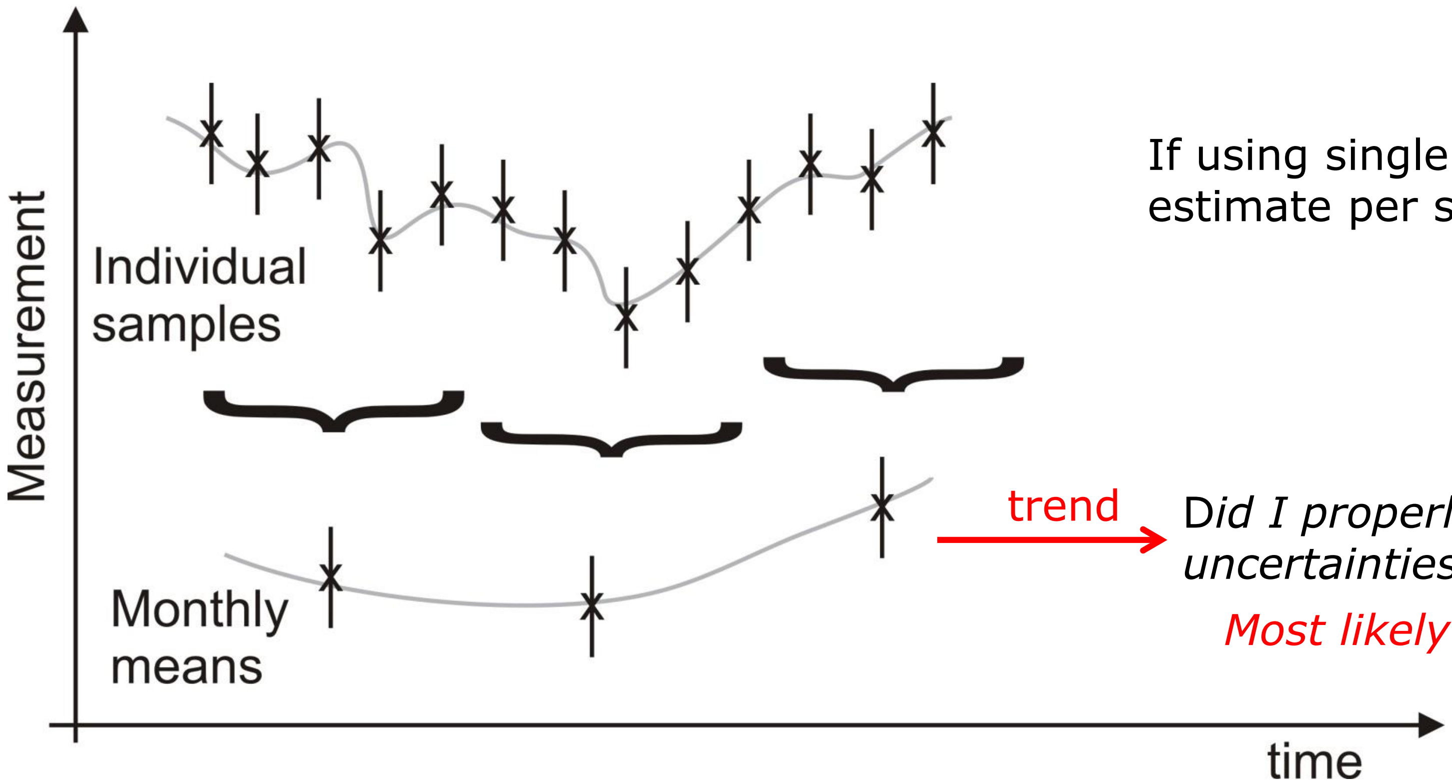
- *NDACC Lidar PIs must document correlation properties and dependencies of input quantities and their associated uncertainty*
- *NDACC PIs must document correlation properties and dependencies of all ozone uncertainty components (systematic vs. random?, in what dimensions? Etc.)*

➔ *Critical for proper handling of "Level-3+ data" (climatologies, assimilation in models, trend studies, etc.)*



Did I properly propagate my uncertainties here?

Use of Uncertainty Information by NDACC Data Users

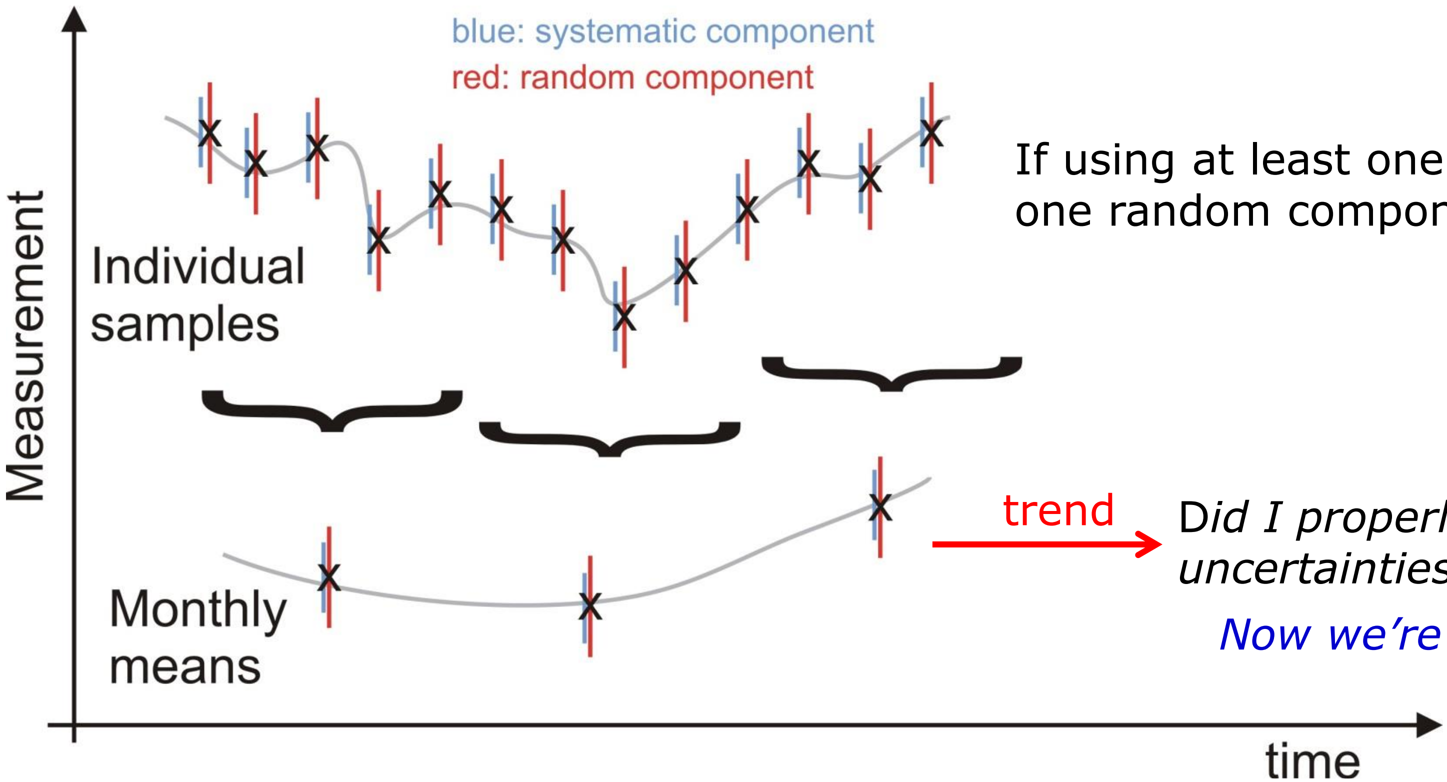


If using single uncertainty estimate per sample

trend → *Did I properly propagate my uncertainties here?*

Most likely NOT

Use of Uncertainty Information by NDACC Data Users



If using at least one systematic and one random component per sample

Did I properly propagate my uncertainties here?

Now we're talking….

Most trend analysis techniques can analytically handle the propagation of sophisticated expression of uncertainty ➔ Let's not be shy about it

Above statement applies likewise for single or combined datasets

CONCLUSION

For the first time in 20 years, NDACC Lidar Group went through a major redesign of their metadata definitions

Full impact of those changes will not be seen until all NDACC lidar datasets are fully re-analyzed using new definitions and approaches

There is a plan to extend this work to Water Vapor and Aerosol Lidars (new ISSI Team?)

More details on this work is available in 3 companion papers:

Leblanc, T., et al.: Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms

– Part 1: Vertical resolution, *Atmos. Meas. Tech.*, **9**, 4029-4049, doi:10.5194/amt-9-4029-2016, 2016

– Part 2: Ozone DIAL uncertainty budget, *Atmos. Meas. Tech.*, **9**, 4051-4078, 10.5194/amt-9-4051-2016, 2016

– Part 3: Temperature uncertainty budget, *Atmos. Meas. Tech.*, **9**, 4079-4101, 10.5194/amt-9-4079-2016, 2016

..and in the "ISSI Team Report" (soon to be WMO Tech. Report):

http://www.issibern.ch/teams/ndacc/ISSI_Team_Report.htm

INTERNATIONAL SPACE SCIENCE INSTITUTE

THANK YOU

BACKUP SLIDES

Motivation

Why a standardized definition of vertical resolution?

- *Ozone DIAL raw data typically needs some smoothing at some point*
- *Actual resolution of the instrument degraded by vertical filtering during data processing*

Simple formulation of the smoothing process:

$$S_f(k) = \sum_{n=-N}^{N} c_n S(k+n)$$

➔ *Instrument "vertical sampling resolution"*

degraded to

ozone profile "vertical resolution"

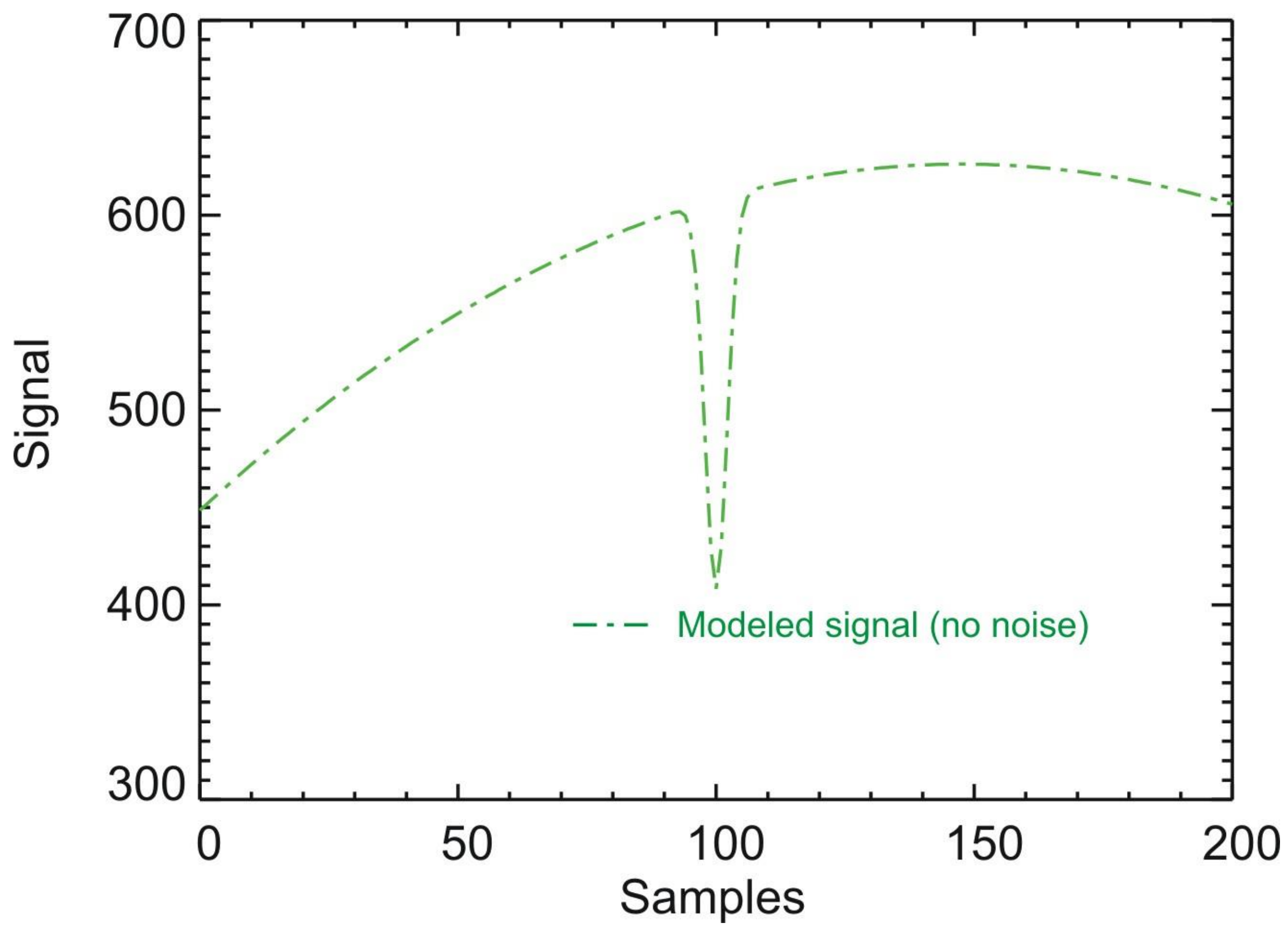
Motivation

There are various ways to report vertical resolution in data files

Example highlighting ambiguity:

We start with a modeled signal

➔ green curve



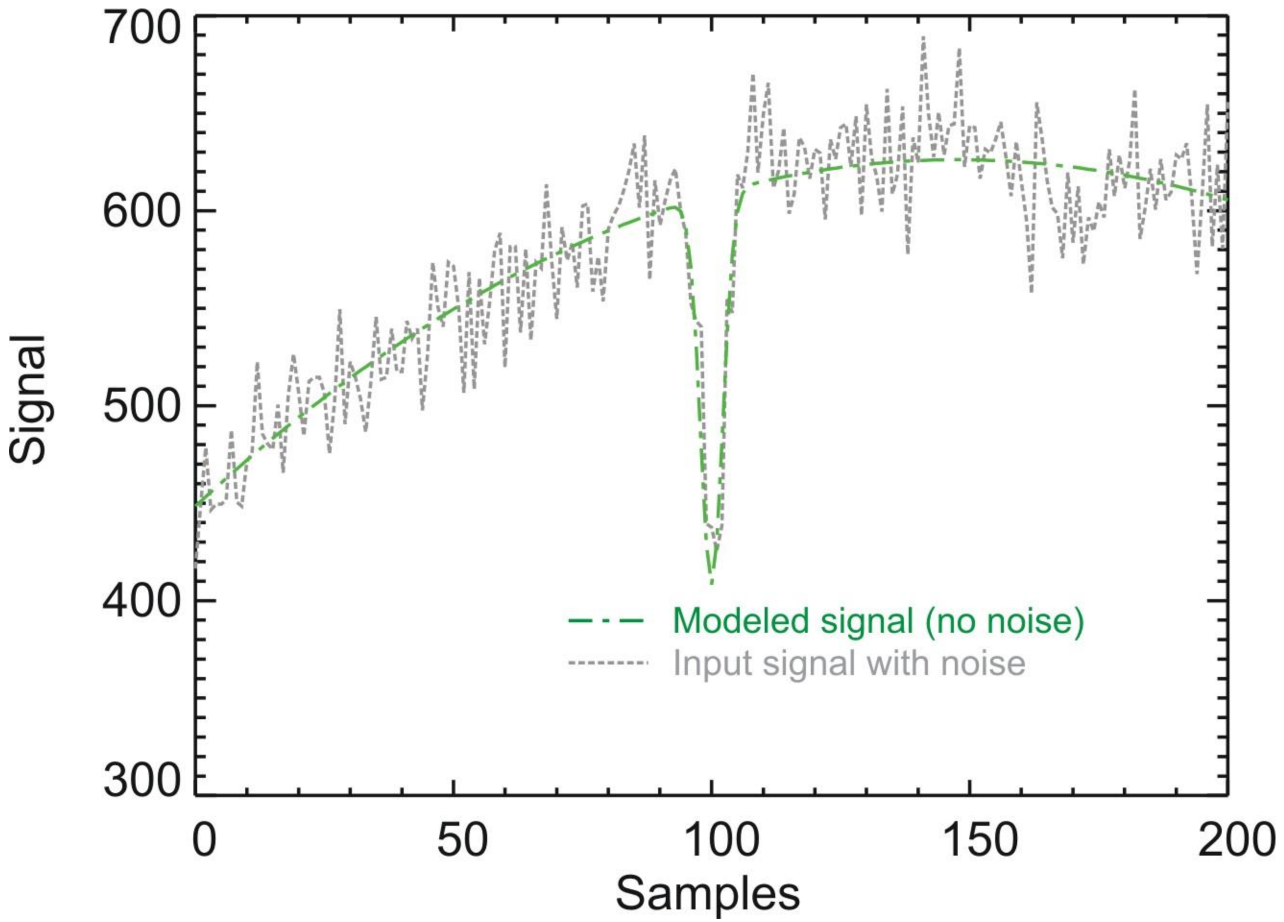
Motivation

There are various ways to report vertical resolution in data files

Example highlighting ambiguity:

We add noise to make it look like a real signal

➔ grey curve



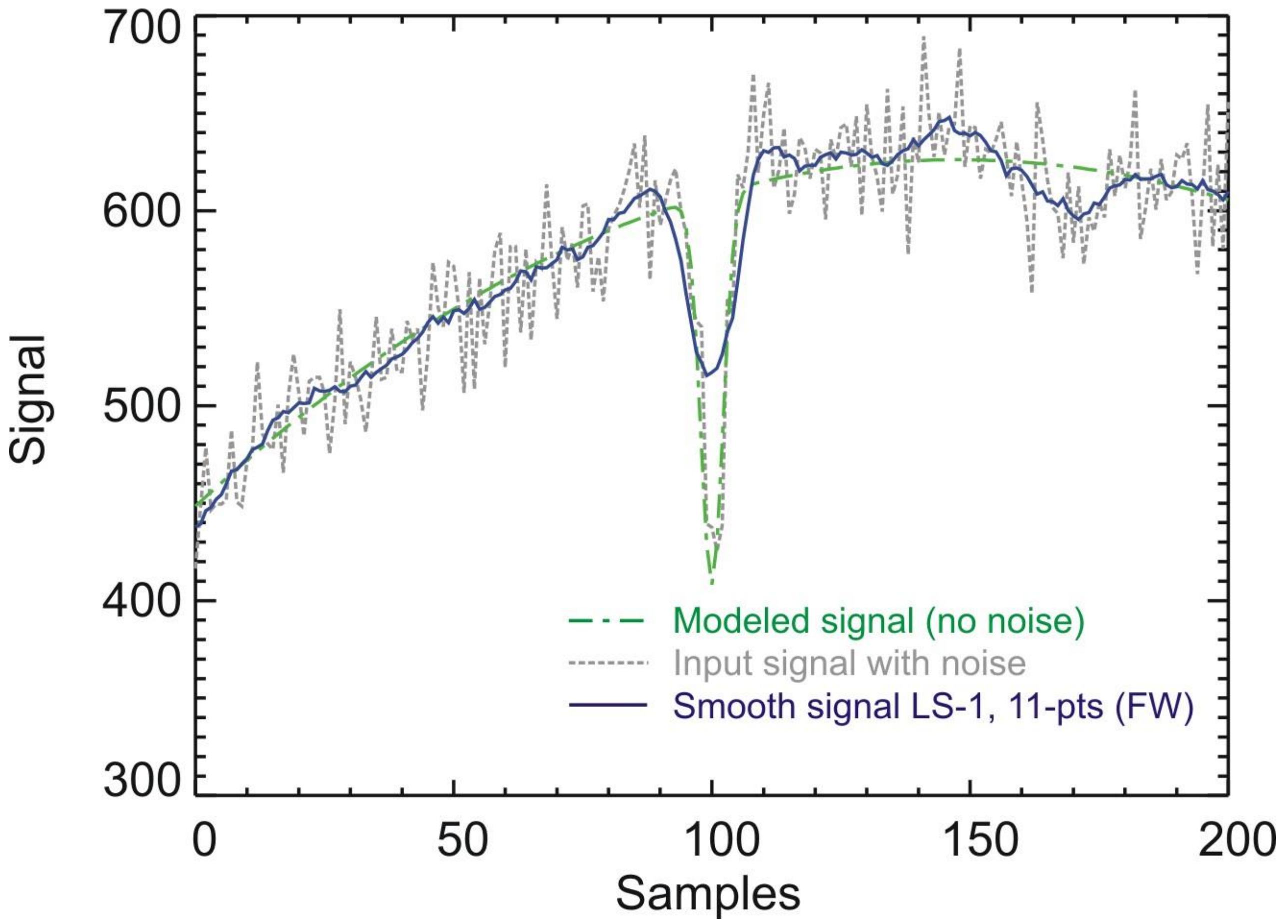
Motivation

There are various ways to report vertical resolution in data files

Example highlighting ambiguity:

We then smooth it with 11-pts FWHM linear fit

➔ blue curve



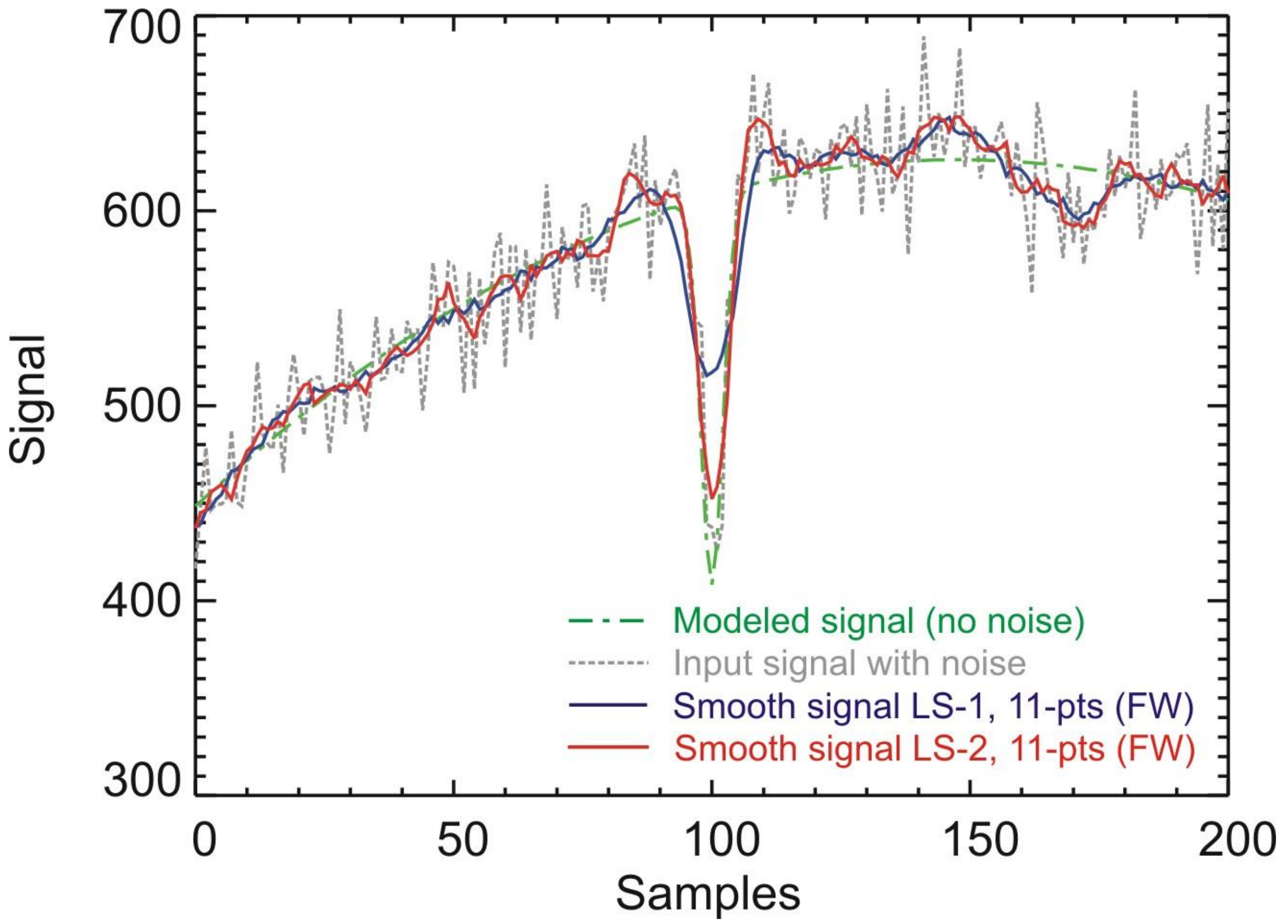
Motivation

There are various ways to report vertical resolution in data files

Example highlighting ambiguity:

We also smooth it with 11-pts FWHM polynomial fit degree-2

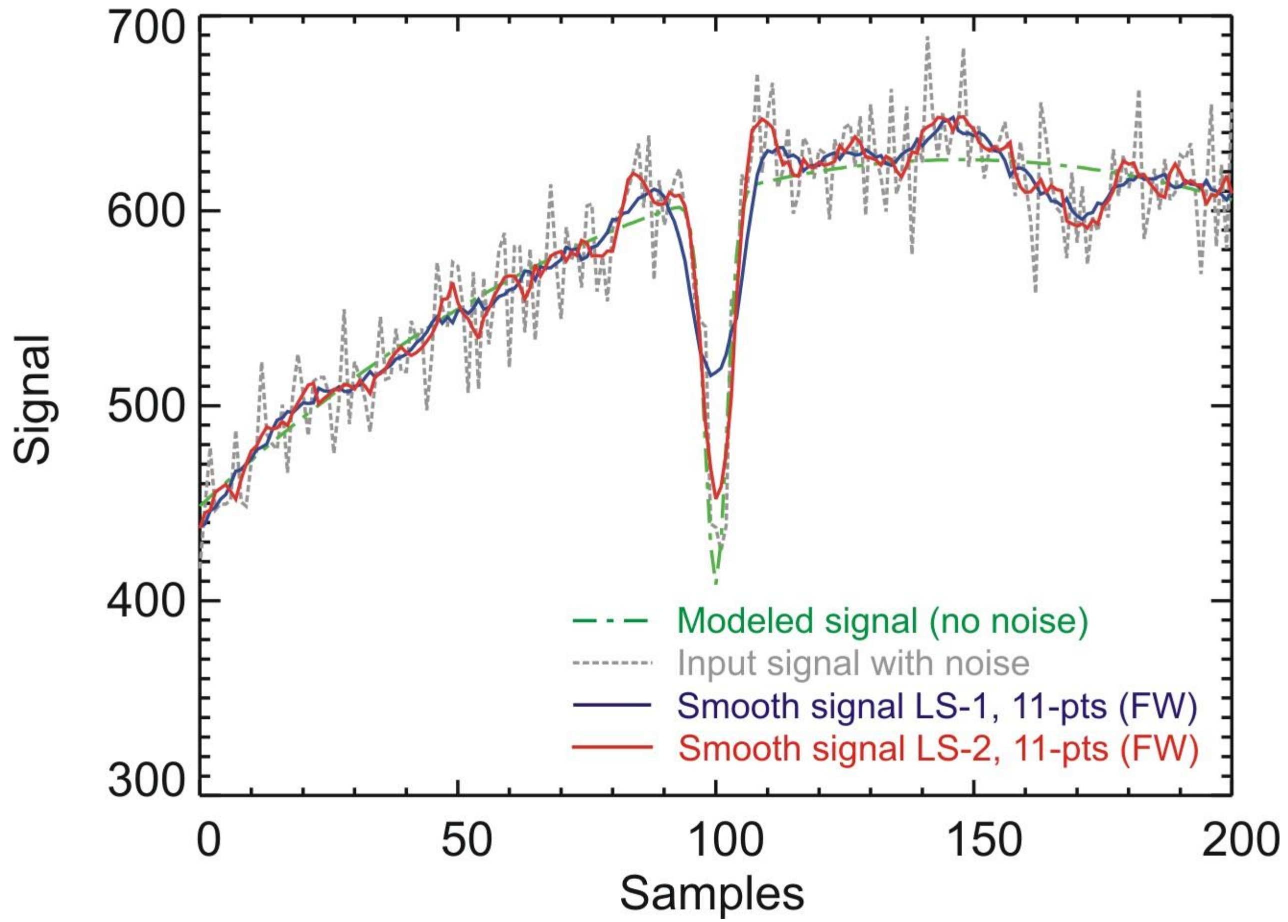
➔ Red curve



Motivation

***Problem**: With the same number of filter coefficients (11 points in this example), we obtain different answers ➔ reporting vertical resolution is ambiguous*

***Solution**: Find a "unique" definition that works well for most NDACC Lidar PIs*



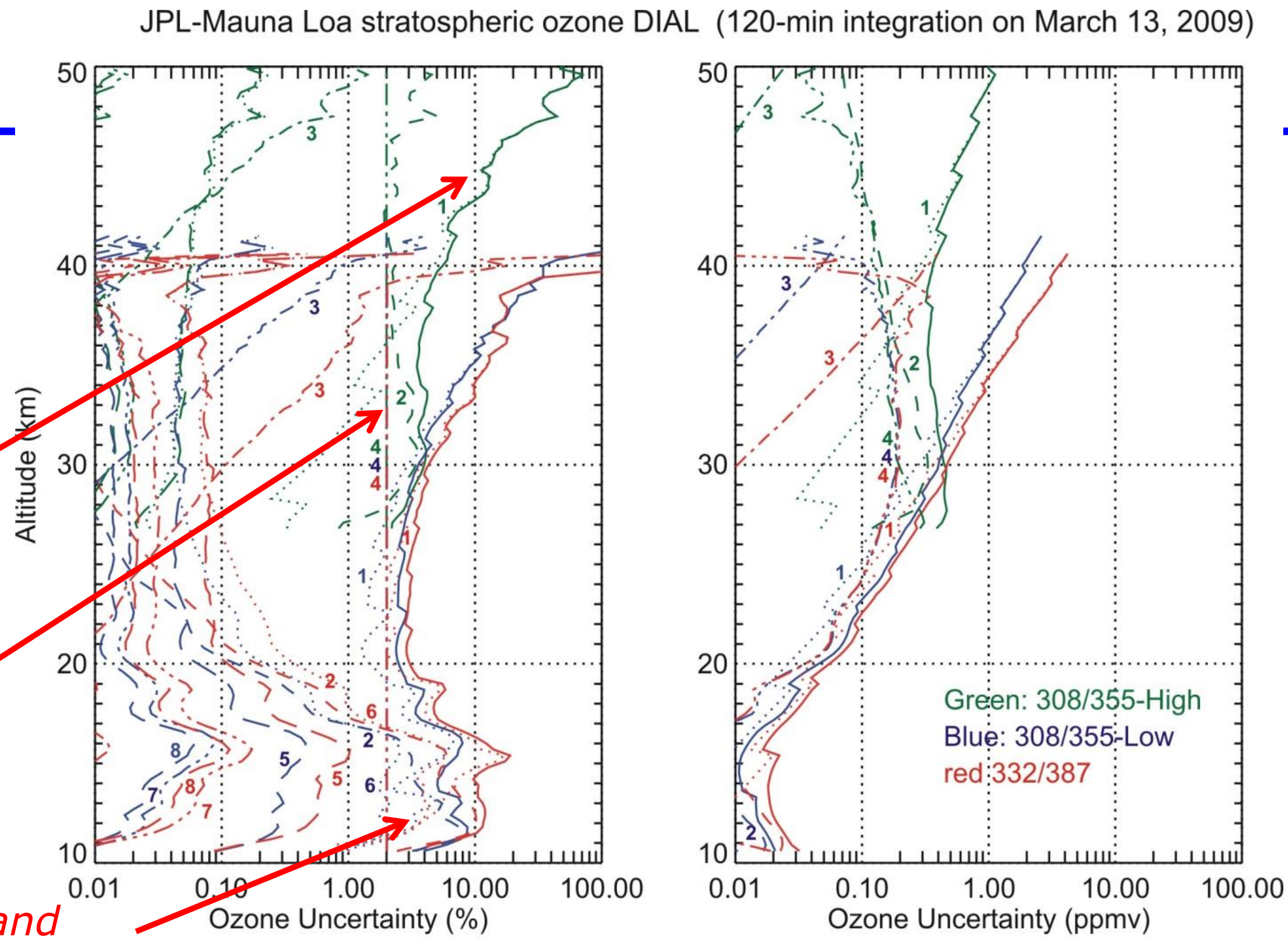
Quantitative example 1:

JPL stratospheric ozone lidar at Mauna Loa, Hawaii

Detection noise uncert. becomes dominant at 38 km (4% and up)

2% "incompressible" uncert. owed to O_3 absorp. cross-sect. whole stratosphere

Saturation correction and molecular extinction cross-sect. largest components below 20 km



- Combined standard uncertainty
- 1 Contribution from detection noise
- 2 Contribution from saturation
- 3 Contribution from background noise
- 4 Contribution from ozone absorption cross-sections
- 5 Contribution from a priori air number density
- 6 Contribution from Rayleigh cross-sections
- 7 Contribution from a priori NO2 number density
- 8 Contribution from NO2 cross-sections

NOTE:

All NDACC lidars are different and there are as many different uncertainty budgets as instruments